

University of Groningen

Acting against one's best judgement

Peijnenburg, Adriana Johanna Maria

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

1996

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Peijnenburg, A. J. M. (1996). *Acting against one's best judgement*. [Thesis fully internal (DIV), University of Groningen]. s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

PART TWO

REVIVAL IN CONTEMPORARY PHILOSOPHY

CHAPTER II

HEMPEL'S ARGUMENT ON ACTION EXPLANATION

0. Introduction

Sometimes people fail to find a solution for a particular problem, and by their very failure solve a completely different problem. The phenomenon is called serendipity and the history of science contains many examples of it. Alfred Nobel failed in his attempt to stabilize nitroglycerine, and thus discovered dynamite; Clyde W. Tombaugh found Pluto near an already predicted position, but the discovery was a fluke because Pluto's low mass could not have caused the apparent perturbations; Arno Penzias and Robert Wilson tried to study local galactic signals, but instead discovered the 'relic radiation' of the Big Bang; Friedrich Wöhler failed to produce ammonium cyanate, but instead synthesised urea and founded organic chemistry; and in failing to find an alternative route to the Orient, Columbus discovered a New World.

The way in which philosophers of the twentieth century have stumbled upon the ancient problem of *akrasia* looks like the exact reverse of a case of serendipity. For here it is not the failure to solve a particular problem that induces the solution to another problem. On the contrary, it is the invention of a dignified solution for a problem that brings another, unsolved and almost forgotten problem back to life. The unresolved and nearly forgotten problem is the problem on the possibility of akratic actions. The dignified solution that caused its revival is Hempel's answer to the question: 'What are the function and the characteristics of a scientific explanation?'

Hempel introduced his theory of scientific explanation in the early forties, but similar views had already been put forward by Popper (Popper 1934); of course, the ideas themselves had been anticipated in the first half of the nineteenth century by Comte's *Cours de philosophie positive* and Mill's *A System of Logic*. Bearing in mind that it grew out of discussions with Paul Oppenheim, I shall mostly refer to Hempel's theory as the H(empel)-O(ppenheim) model; now and then I will also use the names 'H(ypothetico)-D(eductive) theory' or 'covering law model'.

II: Hempel's Argument on Action Explanation

As is well known, the H-O model distinguishes three categories of explanation: deductive-nomological, inductive-statistical, and deductive-statistical explanations. Under the headings of these categories a whole range of explanations fan out: causal, dispositional, functional, teleological, genetic, co-existensive, elliptical, complete, incomplete, and partial explanations are all part of the H-O model. Depending on the particular science as well as on the phenomenon under consideration, the one explanation will of course be more appropriate than the other. However, the essence and also the splendour of the H-O model is that it brings all those different scientific explanations together under a common denominator, consisting of four conditions of adequacy (to be mentioned below). These conditions are common to explanations of miscellaneous phenomena (events, laws, characteristics, trends, animal behaviour, human actions) in miscellaneous sciences (history, biology, physics, sociology, chemistry, psychology). The H-O model thus implies that there is no essential difference between a physical or a chemical explanation of an event, and a psychological or a historical explanation of a human action. From a logical or methodological point of view, the explanation of actions and that of events are supposed to be completely on a par.

Immediately after its publication, the H-O model provoked a passionate and, as it turned out, unabated discussion. Numerous signs of approval and admiration greeted the model, but much disapproval and disagreement has been expressed as well. Several parts of the model were subject to fire, but the part that undoubtedly got the severest attacks involves the claim just mentioned, viz. that there is no logical difference between the explanation of events and that of actions. Many different philosophers have raised many different arguments against that claim. They maintained that the two types of explanation do differ, since the explanation of events may suit the covering law model, but the explanation of actions certainly does not.

In the chapters that make up Part Two I give a general overview of the controversy. My thesis is that the debate, as it progresses, gradually transforms. From being purely a debate on the scope and adequacy of Hempel's model, it little by little becomes a debate on the possibility of *akrasia*. The question of whether the H-O model covers the explanation of actions gradually loses its urgency, whereas the question of whether akratic actions can occur becomes more and more pressing. Part Two contains three chapters, each of which deals with a specific position in the debate. The present chapter is mainly devoted to a description of Hempel's position.

Chapters III and IV are on two important criticisms of Hempel's view: Chapter III deals with those philosophers who took their inspiration from the later Wittgenstein, Chapter IV is about Donald Davidson's critique.

The end of Chapter IV also marks the end of Part Two. By then I shall have explained how the modern debate on action explanation came across the ancient problem of *akrasia*, and how, as a consequence, the debate changed in character, having no longer the correctness of the H-O model, but the possibility of akratic actions as its main issue. Also, I hope to have shown that the *akrasia* problem can never be solved in terms of the debate as it stands. For each of the three debating parties finally runs into the very difficulty. Hence an actual solution of the *akrasia* problem seems to require one of two procedures: either we abandon the present debate and chose a completely new approach, or we retain it and try to repair it from the inside. My own strategy, to be explained in Part Three, is in some sense a compromise between those options. On the one hand I shall step out of the debate and formulate a new approach. On the other, I keep in intimate touch with one position in the debate, namely that of Hempel. In Part Three I shall reshape Hempel's view such that it will be open to degrees, thus making it less classificatory (allowing only yes or no) and more comparative (permitting also gradations). Naturally, it remains to be seen in what sense my strategy can help us in solving "this notorious mess of a problem", as Hintikka called the *akrasia* question (Hintikka 1978).

I start the present chapter by paraphrasing what initially was the key question in the modern debate on action explanation (Section 1). Section 2 and Section 3 together contain a description of Hempel's position. In Section 4 I shortly dwell upon a minor criticism of the Hempelian view. This criticism reflects the arguments that have been raised by the *Verstehen* supporters in the well-known *Erklären-Verstehen* controversy. These arguments used to have, and in a modified form still have, some fame. Nonetheless I call this criticism minor because it is not particularly important for my approach to *akrasia*. The major criticisms of Hempel I discuss in Chapters III and IV.

1. The key question

As is well known, Hempel considers all scientific explanations to be answers to a certain type of why-question, viz. *explanation-seeking why-questions*.

II: Hempel's Argument on Action Explanation

Questions of this type can be expressed in the form 'Why is it the case that p ?', where ' p ' denotes an empirical statement specifying the subject matter of the explanation. 'Why did the television apparatus on Ranger IV fail?', 'Why did Hitler go to war against Russia?' are among Hempel's examples of explanation-seeking why-questions (Hempel 1965, 334). Explanation-seeking why-questions should not be confused with another type of why-question, viz. *reason-seeking* or *epistemic* why-questions. The latter can be expressed in the form 'Why should it be believed that p ?', i.e. 'What reasons are there for believing that p ?' (Hempel 1965). Examples of reason-seeking why-questions are: 'Why would tomorrow's wind-force be 8?' (i.e. 'What reasons are there for believing that tomorrow's wind-force will be 8?') and 'Why would it be that Willard Van Orman Quine derives inspiration from Burrhus Frederic Skinner?' (i.e. 'What reasons are there for believing that Quine is inspired by Skinner?').

Reason-seeking why-questions solicit grounds that will make empirical statements *credible*, whereas explanation-seeking why-questions solicit information that will explain empirical phenomena and thus render them *intelligible* (Hempel 1965, 487-488). The main difference between the two types is that an explanation-seeking why-question normally presupposes the truth of the statement denoted by ' p ', whereas an epistemic why-question does not assume the truth of the corresponding statement, since it requests reasons for believing it to be true. Accordingly, any adequate answer to an explanation-seeking question 'Why is it the case that p ?' must also provide a potential answer to the corresponding epistemic question 'What reasons are there for believing that p ?' (Hempel 1965, 335). Remarkably enough, the converse does not hold. We may have good reasons for believing that tomorrow the wind-force is going to be 8, for instance because the weather forecaster has told us that it will be, but if tomorrow the wind-force is in fact 8, this cannot be explained by pointing to the weather forecast. In other words, the latter is a *reason* for believing that the wind-force is 8, it is not a *cause* for the wind-force being 8 (Hempel 1965, 368).

In Hempel's opinion, answers to explanation-seeking why-questions can all be reduced to a common denominator. This common denominator is the heart of the H-O model; it involves the four conditions of adequacy (CA1)-(CA4):

- (CA1) the statement describing the phenomenon to be explained (this statement is called 'the

- explanandum') must be logically deducible from the information contained in the class of those statements which are adduced to account for the phenomenon (this class is called 'the explanans');
- (CA2) the explanans must contain at least one statement which expresses a general law, and which is required for the derivation of the explanandum;
- (CA3) the explanans must have empirical content; i.e. it must be capable, at least in principle, of test by experiment or observation;
- (CA4) the statements constituting the explanans must be true.¹⁶

Hempel does not consider it a necessary condition for a sound explanation that the explanans contains at least one statement which is not a law. Nonetheless such a statement often appears in the explanans. Taking that into account, the conditions (CA1)-(CA4) yield the following pattern of a scientific explanation, (SE) (Hempel 1965, 249):

(SE)

L_1, L_2, \dots, L_m	general laws
C_1, C_2, \dots, C_n	statements of antecedent conditions
————— [logical deduction]	
E	description of the empirical phenomenon to be explained.

Together, the premises form the explanans whereas the conclusion constitutes the explanandum. Hempel explicitly uses the term 'explanandum' in two ways: "the word 'explanandum' alone will be used to refer either to

¹⁶ Hempel 1965, 247-248. For my present purposes it suffices to formulate (CA1)-(CA4) in the loose and intuitive manner that Hempel chooses (Hempel 1965, 247ff); I have paraphrased that formulation here. We do not need the more precise restatements, see Hempel 1965, 264-278, which after all proved to be false since (i) they countenance the explanation of virtually any particular fact by virtually any theory and (ii) they permit the explanation of any particular fact by itself. (Cf. Hempel's Postscript in Hempel 1965, p. 293ff.)

II: Hempel's Argument on Action Explanation

the explanandum-phenomenon or to the explanandum-sentence: the context will show which is meant" (Hempel 1965, 336). I adopt this double usage and let the context say whether 'explanandum' has the former (ontological) or the latter (linguistic) sense.

The textbook example of an explanation meeting the pattern (SE) is the causal explanation of an event. Using predicate logic, the causal explanation of an event can be represented as (CE):

(CE)

(II.1): $\forall(x) \{P(x) \rightarrow Q(x)\}$	general causal law
(II.2): $P(a)$	statement of antecedent condition
(II.3): $P(a) \rightarrow Q(a)$	universal instantiation, (II.1)
<hr/>	
(II.4): $Q(a)$	description of the event to be explained

where 'x' ranges over events, ' $P(a)$ ' describes the cause, and ' $Q(a)$ ' describes the effect to be explained. Hence, (CE) is (SE) using a causal law.

The key question in the debate on action explanation is: do explanations of human actions meet the four conditions (CA1)-(CA4)? In other words, is the basic pattern of action explanation similar to (CE)? According to Hempel it is. Section 2 and Section 3 offer a précis of Hempel's argument. They summarise Hempel's general attitude towards the relation between action explanations and causal explanations. It is shown that the concept of a disposition is, so to speak, the medium by which Hempel reduces the former to the latter.

2. Action explanations as dispositional explanations

Most people would be willing to assert that all actions are events, but not that all events are actions. This suggests that there exist certain conditions which an event must meet in order to become an action. Thus we might ask what must be added to my ocular tic so that it becomes a meaningful wink - with a nod to Wittgenstein's famous example (*Philosophical Investigations*

I, Section 621; cf. The Brown Book, Chapter II, Section 11-14). But of course the suggestion is misleading: nothing is added to my tic so that it becomes a wink, for an action is not an event plus something else. What distinguishes actions from events are the conceptual frameworks in which they feature. As Wittgenstein might have said: actions and events are flaunted in different language games (*Philosophical Investigations I*, Sections 630-633). I would prefer to say: actions and events are *explained* differently.

The standard way of explaining an event is by pointing out its causes, while the standard way of explaining an action is by reference to its reasons. Thus, when we ask 'Why did the power fail?' we ask what *caused* the failure of the power, but when we ask 'Why did John turn off the main switch?' we ask what John's *reason* was for performing his action.

Hempel obviously is in the same mind about this point: in the H-O model events are explained by causes whereas actions are explained by reasons. Now it is clear enough what Hempel means by the term 'causes', namely events that are described as antecedent or initial conditions in an explanation of the form (CE):

"A 'cause' must be allowed to be a more or less complex set of circumstances and events, which might be described by a set of statements C_1, C_2, \dots, C_k Thus the causal explanation implicitly claims that there are general laws - let us say, L_1, L_2, \dots, L_r - in virtue of which the occurrence of the causal antecedents mentioned in C_1, C_2, \dots, C_k is a sufficient condition for the occurrence of the explanandum event." (Hempel 1965, 348-349).¹⁷

¹⁷ Hempel's notion that causes are events is widely accepted by philosophers of science. Only a few dissidents have actually contradicted it, among whom is John Pollock:

"... in analyzing the concept of a cause philosophers have repeatedly found themselves talking about conjunctions and disjunctions of events. But the relations of conjunction and disjunction are relations between sentences or propositions. If these relations are to be used in some derivative sense in connection with events, this derivative sense must be explained, and such explanation will very likely presuppose facts about the individuation of events

...

However, I think there is a fundamental confusion here. Why do we believe

II: Hempel's Argument on Action Explanation

However, Hempel's use of the term 'reasons' is less clear, because Hempel uses the term in two senses. The first we have already encountered; this sense is present when Hempel contrasts explanation-seeking why-questions with reason-seeking why-questions. Here 'reason' means *reason for believing that something is the case*. The other sense turns up when Hempel contrasts causes of events with reasons for actions. Here 'reason' means *reason for performing an action*. The former sense I shall refer to by reason_{bel}; the latter I describe as reason_{act}. Although Hempel does not explicitly distinguish the two senses, it is obvious from his writings that they should be so separated.

When Hempel is talking about reasons by which actions are explained, he has of course in mind reasons_{act}, not reasons_{bel}. What Hempel has to show, then, is that explanations of actions by reasons_{act} have a structure which is similar to the standard example of (SE), namely (CE). In other words, Hempel has to prove that: "[t]he explanation of an action ..., if adequately formulated, conforms to the conditions for causal explanation" (Hempel 1965, 255).

At this juncture, Hempel can follow either of two lines. First, he can try to show that reasons_{act} are in fact causes, and then claim that explanations by reasons_{act} indeed have the same structure as (CE). Second, he can claim that explanations by reasons_{act} have a (CE)-like structure, although reasons_{act} are not causes. In fact, Hempel chooses the second line. Taking Gilbert Ryle's lead, Hempel regards reasons_{act} as *dispositions*, and explanations by reasons_{act} as *dispositional explanations*. The familiar method of explaining human actions, Hempel declares, is basically of the dispositional kind. To ascribe to an actor reasons_{act} is "to assign to him certain more or less complex dispositional characteristics: this has been

that causation is a relation between events? Most philosophers take this as just obvious, but I do not think that it is at all obvious." (Pollock 1976, 146).

Another non-conformist is Sydney Shoemaker. Like Pollock Shoemaker realises that "it is events, rather than objects or properties, that are usually taken by philosophers to be the terms of the causal relationship" (Shoemaker 1974, 206). He is however of the opinion that "when one event causes another, this will be in part because of the properties possessed by their constituent objects" (ibid.). Thus Shoemaker concludes: "any account of causality as a relation between events should involve, in a central way, reference to the properties of the constituent objects of the events" (Shoemaker 1974, 207).

argued in detail by Ryle, whose ideas have had great influence on the discussion of the subject" (Hempel 1965, 457). On the basis of Ryle's example, Hempel formulates the dispositional explanation as (DE):

(DE)

(II.5): The pane was sharply struck by a stone at time t_I

(II.6): The pane was brittle at time t_I

(II.7): Any brittle object, if sharply struck at any time,
breaks at that time

(II.8): The pane broke at t_I

(Hempel 1965, 459). (DE) can be generalised as (DE*):

(DE*)

(II.9): i was in S at t_I

(II.10): i has D at t_I

(II.11): Any y with D will, when in S at any t , behave in manner R
at t

(II.12): i behaves in manner R at t_I ,

where ' S ' is a situation, ' D ' is a disposition, ' R ' is a way of reacting, ' x ' and ' i ' are objects or individuals, and t, t_I indicate the time (Hempel 1965, 462).

Evidently, dispositions differ from causes: the latter are events while the former are inclinations to behave in certain observable ways under certain observable circumstances. Hempel clearly recognises this difference when he states that

"... possession of [a] dispositional property ... would not
ordinarily count as a cause ..." (Hempel 1965, 487).

On the other hand, however, he refuses to designate dispositional explanations as being downright noncausal:

Dispositional explanations ... cannot be said to be

II: Hempel's Argument on Action Explanation

noncausal" (Hempel 1965, 486).

But if dispositions do not count as causes, in what then does the causal character of a dispositional explanation lie? To that question Hempel answers:

"... a dispositional explanation invokes, in addition to the appropriate dispositional property [*D*], also the presence of circumstances ... in which the property [*D*] will manifest itself by the symptom ... to be explained. *Here [the circumstances] may be said ... to have caused the explanandum event*" (Hempel 1965, 486; my emphasis).

In other words, the causal character of a dispositional explanation lies in the conditions under which the disposition manifests itself, c.q. the conditions described in (II.5) and in (II.9). (II.5) and (II.9) describe the causes of the explanandum events in (DE) and (DE*), in much the same way as (II.2) describes the cause of the explanandum event in (CE). Thus Hempel's argument for the claim that action explanations are causal explanations consist of two simple parts: (i) the statement that action explanations are dispositional explanations, and (ii) the claim that dispositional explanations have the same structure as causal explanations.

There is more to the story, however. In Hempel's view, action explanations are not just ordinary dispositional explanations, but they constitute a special kind; they are also *rational* explanations. Although most philosophers know the story about rational explanations, I nonetheless restate its *Leitmotive* below (in Section 3). On the basis of this story, it can then be explained how not only Hempel's ideas, but those of his antagonists too ran into the time-honoured *akrasia* problem.

3. Action explanations as rational explanations

The idea that action explanations are dispositional explanations was already present in Hempel's earliest work on scientific explanation. But it was not until 1957 that he started to work out this idea in more detail. In that year William Dray published *Laws and Explanation in History*, a much-discussed criticism of the H-O model which gave voice to all those historians and

social scientists who felt uncomfortable with Hempel's rational reconstruction of their daily activities.

As is widely known, Dray argued that the H-O model can never cover the ordinary explanation of human actions. The latter explanation, so the criticism runs, is a "rational explanation" which "displays the rationale of what was done" (Dray 1957, 124). This means that such an explanation refers, not to an empirical law, but to a "principle of action", i.e. "a judgement of the form: 'When in a situation of type C_1, \dots, C_n the thing to do is $A-I$ '" (Dray 1957, 132; Dray's text has ' x ' instead of ' $A-I$ '). Since Hempel deems the presence of an empirical law to be a *sine qua non* of every sound explanation, a rational explanation is said to defy the Hempelian covering law model.

In response to this objection Hempel has introduced his schema (AE^h) of action explanation:

(AE^h)

(II.13): In a situation of type C , any rational agent will do $A-I$	general law
(II.14): P was a rational agent	initial condition
(II.15): P was in a situation of type C	initial condition
<hr/>	
(II.16): Therefore, P did $A-I$,	explanandum

(Hempel 1965, 470ff; Hempel's text has ' A ' instead of ' P ' and ' x ' instead of ' $A-I$ '). Here ' $A-I$ ' describes an action, element of the set A of k possible actions, $A-1, A-2, \dots, A-k$; ' C ' abbreviates Dray's ' C_1, C_2, \dots, C_n ' (Hempel 1965, 470; cf. Dray 1957, 132); according to Hempel ' C ' describes three items:

- (a) the end the agent sought to attain;
- (b) the agent's beliefs concerning the empirical circumstances in which he had to act and concerning the means available to him for the attainment of his objective;
- (c) moral, religious, or other norms to which the agent was committed.

Item (b) contains what Hempel calls the *information basis*, i.e. the set of

II: Hempel's Argument on Action Explanation

sentences that the agent believes to be true. Items (a) and (c) jointly constitute what Hempel calls the *total objective*, i.e. the "set *E* of sentences describing the intended end state, in conjunction with a set *N* of constraining norms", where *N* contains "certain general constraining principles, ... moral or legal norms, contractual commitments, social conventions, the rules of the game being played, or the like" (Hempel 1965, 464ff). Thus '*C*' in Hempel's schema denotes the union of the agent's total objective and his total information basis.

(AE^h) is a dispositional explanation, for having a certain objective and possessing a certain amount of information both are dispositions. But (AE^h) is also a *rational* explanation. Before I expand that point in Section 3.2, four comments have to be made.

3.1 Four comments

Hempel's schema for action explanation (AE^h) calls for four comments, most of which are merely terminological. The first is about the make-up of the total objective. It is unclear why and how Hempel divides it into two parts: *E* or (a) on the one hand and *N* or (c) on the other. As far as I know we do not have an adequate theory of differences between norms and intended end states. Moreover, even if we were to have such a theory, it would be doubtful whether the presumed differences would be of any significance to the *explanation* of actions: relative to a theory on action explanation, end states and norms are completely on a par (as would be, for instance, what Kant called *Pflicht* and *Neigung*). But the main reason for calling a distinction between *E* and *N* unclear is that it is illogical in the light of Hempel's own thoughts. According to Hempel, *E* as well as *N* contain dispositions and he nowhere even tries to discriminate between the disposition to pursue an end and the disposition to obey a norm. He regards both dispositions as cases of the general umbrella disposition that other authors have termed a 'pro-attitude' or 'want' or 'desire'. The point is neatly described by Tuomela:

"We still lack a general satisfactory account on the similarities and dissimilarities between the conceptual behavior of different psychological pro-attitudes (such as ... volitions, wants, desires, etc.). In any case I am willing

to submit that they all intrinsically contain a disposition to behave towards a goal, and this in fact suffices for our present discussion" (Tuomela 1976, 204, footnote 7).

Like Tuomela, Davidson bestows upon the term 'pro-attitudes' a very broad meaning. Under pro-attitudes are to be included "tastes, values, principles, inclinations, and other evaluative attitudes" (Davidson 1985c, 202); pro-attitudes contain "desires, wantings, urges, promptings, and a great variety of moral views, aesthetic principles, economic prejudices, social conventions, and public and private goals and values" (Davidson 1963, 4). In sum, then, "[t]he word 'attitude' does yeoman service here, for it must cover not only permanent character traits that show themselves in a lifetime behaviour, ... but also the most passing fancy that prompts a unique action ..." (Davidson 1963, 4). Davidson concludes:

"It is not unnatural, in fact, to treat wanting as a genus including all pro attitudes as species" (Davidson 1963, 6).

Robert Audi is another scholar who has argued that we should use one broad, general motivational notion; and like Davidson, he thinks the concept of wanting can do the job (Audi 1973a; for an explication of this concept, see Audi 1973b). Still another supporter of the view in question is Christopher Peacocke. According to him, any distinction within the class of desires or pro-attitudes is irrelevant to the understanding of akratic actions:

"In recent writing on the philosophy of action several important distinctions have been made between various kinds of desires. Distinctions have been drawn between first- and second-order desires, between desires one desires to be acting on and those which one does not desire to be acting upon, and so forth. But I would claim that almost all these distinctions are irrelevant or only obliquely relevant to the problem of *akrasia*. No intrinsic distinction *within* the class of desires can provide a full understanding of *akrasia*." (Peacocke 1985, 51).

As far as this particular point is concerned, I agree with all the authors mentioned: in compliance with Tuomela, Davidson, Audi, and Peacocke (and

II: Hempel's Argument on Action Explanation

contra for instance Milligan who distinguishes 'wanting' from 'liking' - Milligan 1980, 31), I shall not discriminate between pursuing a goal and obeying a norm. Both of them I consider to be pro-attitudes towards actions of a certain kind. Consequently I neglect any distinction between (a) and (c); for me, intended end states E and constraining norms N are entirely on a par. I will treat E and N as species of a genus, and as generic term I choose *pro-attitude*, sometimes also *want* or *desire*. This generic term I take to denote everything that can be regarded as an attitude of approval towards actions of a certain kind. The m members of $E \cup N$ I refer to as the set D of pro-attitudes or desires: $D = \{D-1, \dots, D-m\}$. The n elements of the information basis will be termed beliefs: $\{B-1, \dots, B-n\} = B$. Accordingly, the 'situation of type C ' in schema (AE^b) refers to the union of B and D , $B \cup D$.¹⁸

¹⁸ David Milligan is not the only author who makes out a case for a distinction between pro-attitudes. Michael Bratman is another. Already in his early writings, Bratman maintains that there is an essential difference between deliberating about *what would be best to do* and deliberating about, say, *what would be chic to do* (Bratman 1979, 158). The former is, as he calls it, "evaluative" in a way that the latter is not. According to Bratman, only evaluative deliberations about what would be best to do lead to "full-blown actions", i.e. actions that are "free, deliberate, purposive" (ibid., 154). The idea that such actions could also result from other deliberations (for instance, deliberations about what would be chic to do) Bratman calls an "extreme externalist response":

"This response finds no essential relation between desiring to do something and valuing doing it. It sees judgements about what would be best as playing no special role in the practical reasoning."
(ibid., 159).

Whatever the merit of Bratman's distinction between desiring and valuing may be, I do not think it is relevant in the present context. That context is, I recall, dominated by the *explanation* of actions. Since an action, A , can be explained equally well by saying that the actor *desired* A or by saying that he *valued* A , the difference between desiring and valuing is of little significance here. Moreover I believe that the difference is not relevant to the *akrasia* problem either. For the latter can equally well be phrased by saying that the agent acts against what he *values*, or acts against what he *desires* most vehemently.

In later writings Bratman confirms that there exists a distinction among pro-attitudes; now the distinction is between desires and intentions: "Both intentions and desires are ... *pro-attitudes*. ... Intentions are, whereas ordinary desires are not, *conduct-controlling* pro-attitudes. Ordinary desires, in contrast, are merely *potential influencers* of

The second comment has to do precisely with this union. I take $B \cup D$ to constitute the set of *reasons* for an action. (I will from now on use the term ‘reason’ whenever I mean ‘reason_{act}’ - cf. Section 2.) In conformity with the current literature on the subject, I assume that a reason for an action is a pair, consisting of a belief and a desire. For example, if ‘ $D-i$ ’ ($1 \leq i \leq m$) means ‘wanting such and such to be the case’ and ‘ $B-j$ ’ ($1 \leq j \leq n$) means ‘believing that performing action $A-h$ ($1 \leq h \leq k$) establishes such and such’, then together $D-i$ and $B-j$ form a reason for doing $A-h$. There exists a weak tendency to use the term ‘reason explanation’ instead of ‘action explanation’. Despite the confusion engendered by this usage (after all a reason is part of the explanans and not of the explanandum), I shall follow this practice and use the terms ‘action explanation’ and ‘reason explanation’ interchangeably.¹⁹

It should be kept in mind - and this is the third comment - that in a reason explanation like Hempel’s schema (AE^h) no requirements are imposed on the beliefs and the pro-attitudes. That is, the beliefs need not be true nor need the desires to be virtuous or honourable. Whenever we explain a person’s action by an appeal to his reasons, we are interested in his factual beliefs, not in the grounds on which he holds them; similarly, we are interested in what the agent in fact desires, not in the moral merit of his wants (Hempel 1965, 464-465; cf. Introduction to Chapter I).

Comment number four: Hempel nowhere even remotely suggests that the extensions of the belief set B and the desire set D are of any significance for the explanation of actions; the structure of (AE^h) is obviously not affected by the number of beliefs and desires that constitute the situation of type C . Therefore nothing keeps us from supposing that P has only one belief and one pro-attitude and thus that B and D are singletons: $D = \{D-1\}$, $B = \{B-1\}$, where $D-1$ and $B-1$ might be read as ‘wanting $A-1$ ’ and ‘believing that $A-2$ is a means for $A-1$ ’. This enables us

action” (Bratman 1987, 15-16). Bratman’s formulation suggests that he has in mind a distinction between first and second order pro-attitudes: I say more about that distinction in the Introduction to Chapter V. Furthermore, the notion of intention has its own problems. I say more on the term ‘intention’ in Chapters III, IV, and VI.

¹⁹ Some philosophers also use the word ‘intentional explanation’. That usage I shall avoid.

to simplify (AE^h) to (AE^{hs}):

(AE^{hs})

(II.17): $\forall(x) \{ (R(x) \wedge D-I(x) \wedge B-I(x)) \rightarrow A-I(x) \}$

(II.18): $R(a)$

(II.19): $D-I(a)$

(II.20): $B-I(a)$

(II.21): $A-I(a)$.

Sentence (II.17) replaces (II.13); (II.18) is a translation of (II.14), and for (II.16) sentence (II.21) is substituted. Together, (II.19) and (II.20) replace (II.15); (II.19) says that person *P* has pro-attitude *D-I* and (II.20) says that *P* has belief *B-I*. In conformity with the conditions of adequacy, it is assumed that (II.17) is a general law, that (II.17)-(II.21) all have empirical content, and that they all are true. In Chapter III we will see that some of these assumptions have come under heavy criticism.

3.2 Acting rationally:

the higher-order disposition to comply with decision theory

In the venerable tradition of Peirce, who claimed that "different beliefs are distinguished by the different modes of action to which they give rise", Hempel considers beliefs and desires as entities that can only be distinguished on the basis of actions (Peirce, *Collected Papers*, vol. 5, Section 398ff). Hence beliefs and desires are dispositions, and (AE^{hs}) is a dispositional explanation. It basically has the same structure as (DE*), which in turn resembles (CE), the standard example of the scientific explanation (SE).

However, (AE^{hs}) is also a *rational* explanation. Explaining *P*'s action not only requires that we formulate a full statement of *P*'s desires and beliefs, it also demands a demonstration to the effect "that the action was to be expected in view of those objectives and beliefs" (Hempel 1965, 469). By this Hempel means that the reasons involved have to be *good reasons*. This is of course not to say that the beliefs have to be true or that the objectives have to be noble. Rather, it is to say that *P*'s action "suits" *P*'s reasons

(Hempel 1965, 465); as judged in the light of *P*'s objectives and beliefs, his action appears to be "*reasonable* or *appropriate*" (Hempel 1965, 463; emphasis by Hempel). Hence we have to attribute to *P* another trait, viz. *R*, the property of being a reasonable or rational agent. This property is assigned in the so-called *rationality assumption*: sentence (II.14) in (AE^h) and (II.18) in (AE^{hs}).

Like having a belief or a desire, being a rational agent is a disposition; it is a tendency to behave in a characteristic way under circumstances of a given kind. But unlike the other two, the rationality disposition is a *higher-order disposition*. For the particular circumstances under which a rational agent is said to behave in a characteristic way include dispositions, viz. the agent's desires and beliefs. Hence those conditions cannot simply be described in terms that refer to observable, external situations:

"When we call someone a rational agent, we assert by implication that he will behave in characteristic ways if he finds himself in certain kinds of situation; but such situations cannot be described simply in terms of environmental conditions and external stimuli; for characteristically they include the agent's objectives and his relevant beliefs. To mark this difference, we might say that the dispositions implied by attributing rationality to a person are *higher-order-dispositions*; for the beliefs and ends-in-view in response to which, as it were, a rational agent acts in a characteristic way are not manifest external stimuli but rather, in turn, broadly dispositional features of the agent." (Hempel 1965, 473).

Stated in this way, the concept of rationality is not very precise, but thanks to the theory of decision making we are able to make it more specific. Decision theory has supplied us with quantitative terms that can be substituted for the merely qualitative terms that we use when we normally speak about rational and irrational actions. Thus the qualitative terms 'belief' and 'partial belief' can be replaced by the quantitative terms 'subjective probability', and the qualitative 'desire' and 'pro-attitude' can be replaced by the quantitative 'subjective desirability' or 'subjective utility'. Accordingly, the concept of being a rational agent can be given a

quantitative shape too.

Hempel has accepted the gifts of decision theory with appreciation, notably its descriptive (as opposed to its normative) and its individual (as opposed to its game theoretical) variant. How to specify *P*'s higher-order disposition to act rationally, i.e. his ability to make his action at *t* "suit" his reasons at *t*? What exactly does 'suitability' mean in this case? Hempel's answer to this question is couched in terms of descriptive individual decision theory, with its various rules (the rule of maximizing expected utility, the maximin rule, the maximax rule, et cetera) for various cases in which decisions have to be made (decisions under certainty, under risk, under uncertainty, et cetera). Hempel's final answer is that *P*'s action at *t* suits his reasons at *t* if and only if *P* follows a rule of which decision theory says that it applies to the situation *P* is in at *t*.

The view that rational acting equals bowing to decision theoretical dictates might sound rather uncomfortable, to say the least. Even if we do not bother about philosophers like Habermas, but simply restrict rational acting to rude and simple *zweckrationales Handeln*, problems remain. For decision theory, that mathematical model for this conception of rationality, is full of disconcerting dilemmas and puzzling paradoxes.

No matter how threatening and how enchanting the prisoner's dilemma and Newcomb's problem may be, I shall ignore them here. I circumvent the pitfalls of decision theory and accede to Hempel's belief that decision theory is an effective mathematical representation of rational acting. For here, in Part Two, I am interested in another piece of Hempel's analysis. I am interested in the idea of rationality as a *higher-order disposition*, not in the decision theoretical *specification* of that idea. I shall argue, mainly in Chapter III, that this idea shifted the focus of the debate on action explanation and made the old *akrasia* problem a pressing issue again.

4. Criticism of Hempel's argument

4.0 Introduction

On the basis of the above, Hempel's position is easily summarised. Action explanations satisfy the four conditions of a sound scientific explanation, (SE), and hence fully fit into the H-O model. Moreover, action explanations bear a very close resemblance to the textbook cases of the H-O model, viz.

the ordinary causal explanations of events, (CE). For although reasons are dispositions and thus cannot be called causes, the circumstances under which reasons manifest themselves very well might be termed causes. Conclusion: reason explanations resemble causal explanations.

Contrary to Hempel, countless philosophers have argued that explanations of actions differ fundamentally from causal explanations of events. Arguments in defense of this difference have been put forward by philosophers as diverse as R.G. Collingwood, M. Oakeshott, K.O. Apel, G.E.M. Anscombe, W. Dray, P. Winch, S. Hampshire, A.I. Melden, F. Stoutland, J. Habermas, G.H. von Wright, D. Davidson and K. Holzkamp.²⁰ Small wonder, then, that those arguments display a great variety. Although in general it is a risky enterprise to sort arguments into clusters, all the more so if the arguments are gathered from very different philosophical climates, I shall nonetheless run that risk and divide the arguments concerned into two groups. It should be borne in mind, however, that my division simply has a pragmatic aim, namely that of gaining a better understanding of akratic actions. In other contexts, where other aims are at stake, quite different classifications may well be appropriate.

The first group contains arguments which are taken from what became known as the *Erklären-Verstehen* debate. I consider this group as being of minor importance for my purpose, but I nevertheless cast a quick and fairly frivolous glance at it in Section 4.1. The second group of arguments *is* important for my case. It enters when the relation between reasons and causes is discussed. I shall divide this group into two subgroups. The first consists of supporters of the so-called Logical Connection Argument, the second is headed by Donald Davidson. I deal with both groups in Chapter III and Chapter IV respectively.

4.1 The Erklären-Verstehen controversy

The first group contains arguments for the claim that the method used in the social sciences differs from that of the natural sciences, and that, therefore,

²⁰ Collingwood 1946; Oakeshott 1953; Apel 1979; Anscombe 1957; Dray 1957; Winch 1958; Hampshire 1959; Melden 1961; Stoutland 1970; Habermas 1968; Von Wright 1971; Davidson 1963, 1976, 1987; Holzkamp 1972.

II: Hempel's Argument on Action Explanation

action explanations differ from causal explanations. In the present section I give a brief characterisation of those arguments. Then I bypass them in order to fix my attention on the two subclasses of the second group.

Most of the arguments within the first group echo arguments which nineteenth-century hermeneutics might have raised against their positivistic contemporaries. As we all know, the early positivists argued for a unity of scientific method whereby the natural sciences set the standard, dictating that all scientific explanations are subsumptions of individual cases under hypothetically assumed general laws.²¹ In opposition to this methodological monism, hermeneutically *angehauchte* philosophers like Windelband and Dilthey claimed the existence of a method typical of the

²¹ Examples of this viewpoint exist in dozens in the writings of the founders of positivism, Hempel's predecessors Comte and Mill. A random sample choice from the work of Comte (translations by myself):

- "... by *positivistic philosophy* ... I mean merely the proper study of the generalities of the different sciences, conceived as being submitted to a unique method and as forming the different parts of a general plan." ("... par *philosophie positive* ... j'entends seulement l'étude propre des généralités des différentes sciences, conçues comme soumises à une méthode unique et comme formant les différentes parties d'un plan général.")
- The explanation of the facts ... is henceforth no more than a connection established between the various phenomena and some general facts." ("L'explication des faits ... n'est plus désormais que la liaison établie entre les divers phénomènes particuliers et quelques faits généraux.")
- "Thus the real positivistic spirit consists above all in ... studying what is, in order to infer what will be, according to the general dogma of the invariance of the natural laws." ("Ainsi, le véritable esprit positif consiste surtout ... à étudier ce qui est afin d'en conclure ce qui sera, d'après le dogme général de l'invariabilité des lois naturelles".)

Here is the result of a dip into Mill's work:

- "It must be kept constantly in view, therefore, that in science, those who speak of explaining any phenomenon mean (or should mean) pointing out ... some more general phenomenon, of which it is a partial exemplification; or some laws of causation which produce it by their joint or successive action, and from which, therefore, its conditions may be determined deductively";
 - "An individual fact is said to be explained, by pointing out its cause, that is, by stating the law or laws of causation, of which its production is an instance";
 - "An hypothesis is any supposition which we make ... in order to endeavour to deduce from it conclusions in accordance with facts which are known to be real; under the idea that if the conclusions to which the hypothesis leads are known truths, the hypothesis itself either must be, or at least is likely to be, true."
- Quotes could easily be multiplied.

Geisteswissenschaften. The latter have "the historico-social reality as their object" ("die geschichtlich-gesellschaftliche Wirklichkeit zu ihrem Gegenstande", Dilthey 1883, 4), and they think psychology is the "basic science" ("Grundwissenschaft", Dilthey 1895/96, 273). *Geisteswissenschaftler* take their own line, Dilthey and Windelband argue, and this line differs considerably from the procedure followed by natural scientists. A few differences which are frequently mentioned are the following. (All the translations of Dilthey's text are my own.)

To begin with, the *Geisteswissenschaften* have a distinctive *subject matter*. It consists of "facts of the mind" ("geistige Tatsachen"), not as in the natural sciences, of "physical facts" ("physische Tatsachen", Dilthey 1895/96, 248). It affects "internal processes or states" ("innere Vorgängen oder Zuständen"), not as in the natural sciences, "external or perceptual observations" ("äussere oder sinnliche Wahrnehmungen", Dilthey 1895/96, 243-244). It primarily has to do with "the singular" ("das Singulare"), not as in the natural sciences, with "general laws" ("das Gesetzliche", Dilthey 1895/96, 271). The unique nature of the subject matter requires a *specific attitude* on the side of the individual researcher. Contrary to the natural scientist, the individual *Geisteswissenschaftler* has to penetrate into his object with all his heart so as to live and relive it: " ... the *attitude of researchers* in each field is very different ... the art of observing external objects gives the natural researcher a completely different frame of mind than that arising from the historical understanding, the deepening sensitive empathy with human or historical condition." (Dilthey 1895/96, 260; emphasis by the author).²² Accordingly, the application of a *specific method* is required: "What makes the difference already valid with respect to the personal disposition, becomes even more the case so far as the *method* is concerned." (ibid.; emphasis by Dilthey).²³

The main difference between the *Naturwissenschaftliche* and the *Geisteswissenschaftliche* method turns out to be twofold. First, the one

²² " ... der *Habitus des Forschers* auf beiden Gebieten ist gar verschieden die Kunst der Beobachtung äusserer Objekte gibt dem Naturforscher eine ganz andere Geisteshaltung, als welche aus dem geschichtlichen Nacherleben, dem vertieften mit dem Gemüt sich Einleben in menschliche oder historische Zuständlichkeit erwächst."

²³ "Was sich schon als Unterschied in der persönlichen Verfassung geltend macht, das tritt als solcher der *Methoden* noch viel durchgreifender heraus."

II: Hempel's Argument on Action Explanation

makes use of experiments while the other does not: "In the natural sciences the experimental methods have an overriding significance ... but no experimental work has so far led to knowledge of laws in the internal mental domain." (Dilthey 1895/96, 262).²⁴ Second, mathematics is at the heart of the one, while it is only of minor importance to the other: "In a similar way the utilization of mathematics has played a minor rôle in the field of the mental sciences. On the other hand in the mental sciences description (narration), analysis and comparative methods preponderate, whereas in the natural sciences they give way to induction, experiment and mathematical theory." (Dilthey 1895/96, 262).²⁵

All these nineteenth-century arguments have been zealously used by twentieth-century philosophers who aimed to show that the social sciences have their own method, and that therefore explanations of actions escape the straitjacket of the H-O model. Habermas, Holzkamp, Schutz and a host of others have shown themselves veritable Dilthey's *après la lettre* in arguing that the social sciences are unable to perform experiments in which the relevant variables are kept under control. Just like the old *Geisteswissenschaften*, the social sciences have a peculiar subject matter. This subject matter is then explained in phrases which could have been Dilthey's: it is inexhaustibly unique and infinitely complex, it occurs at a given historical period and within a special cultural setting, it changes as a result of being studied, it has essentially value-impregnated aspects, it can itself be the result of social scientific research, et cetera. Naturally, these allegations did not pass unchallenged. Ernest Nagel is among the philosophers who tried to conceive a neo-positivistic counter to this neo-hermeneutic critique, thus giving fresh food for what had become generally known as the *Erklären-Verstehen* debate (Nagel 1961, Chapters XIII and XIV).

In its original form, the controversy about *Erklären versus*

²⁴ "In der Naturwissenschaften besitzen die experimentellen methoden eine herrschende Bedeutung ... aber keine experimentelle Arbeit hat bisher auf innerpsychischem gebiet zur Erkenntnis von Gesetzen geführt."

²⁵ "Ebenso untergeordnet ist die Verwertung der Mathematik auf dem Gebiet der Geisteswissenschaften. Dagegen überwiegen in den Geisteswissenschaften Beschreibung (Erzählung), Analysis und vergleichende Methode, welche in den Naturwissenschaften gegen Induktion, Experiment und mathematische Theorie mehr zurücktreten."

Verstehen debate has never been really resolved. After having had the contents of several philosophical journals in its grasp for a decade or two, it gradually quitted the philosophical scene and finally petered out somewhere in the late seventies. For some years nobody seemed to have any interest in what neo-hermeneutics and neo-positivists had to say to one another. However, the *Erklären-Verstehen* debate appears to be enjoying a revival; the distinction between natural and social science is again a matter of dispute, albeit in a different way. I am referring to the discussion about Giddens notion of 'the double hermeneutic' (Giddens 1976). By 'the double hermeneutic' Giddens means the phenomenon that concepts originally used by social scientists become part of the vocabulary that is spoken by laymen. The social facts which are described/explained by social scientists appear to be described/explained by non-professionals too. As a result, social scientists are engaged in double heremeneutic: they must interpret their subject matter on a scientific *and* on a common sense level. According to Giddens, this two-fold hermeneutic task is peculiar to the social sciences. For although natural scientists certainly interpret the facts that they describe and explain, those facts are not interpreting instances themselves.

The notion of a double hermeneutic aroused an agitated discussion, the end of which does not yet seem to be in sight (cf. *Social Epistemology* 7 (1993), 183-211). Still the hermeneutic, whether single or double, will not concern me here. Any attempt to settle the *Erklären-Verstehen* debate, whether new or obsolete, is beyond my scope. In my view, the arguments taken from this debate seldom affect the finesses of explanation patterns. More importantly, they have no direct bearing on the problem of *akrasia*. Any relation between the *Erklären-Verstehen* debate and the problem of *akrasia* must surely proceed through the set of arguments discussed in the next two chapters.